JOURNAL OF MULTIVARIATE ANALYSIS 39, 79–86 (1991)

Gaussian Conditional Structure of the Second Order and the Kagan Classification of Multivariate Distributions

JACEK WESOŁOWSKI

Warsaw University of Technology, Warsaw, Poland

Communicated by C. R. Rao

The only Kagan-$\mathscr{D}_{n,2}$(loc) probability measure with Gaussian conditional structure of the second order (GCS_2) is a multivariate normal distribution. Also a short review of the known results on the GCS_2 distributions is given. © 1991 Academic Press, Inc.

1. Introduction with a Review of the Known Results

We begin with the following

DEFINITION 1. A real square integrable random element $X=(X_\alpha)_{\alpha\in\mathscr{A}}$ has Gaussian conditional structure of the second order ($X\in \mathrm{GCS}_2(\mathscr{A})$) if for any $\alpha_1, ..., \alpha_n \in \mathscr{A}$, $n=2, 3, ...,$

(i) random variables $X_{\alpha_1}, ..., X_{\alpha_n}$ are linearly independent and pair wisely correlated;

(ii) $E(X_{\alpha_1} \mid X_{\alpha_2}, ..., X_{\alpha_n})$ is a linear function of $X_{\alpha_2}, ..., X_{\alpha_n}$;

(iii) $\mathrm{Var}(X_{\alpha_1} \mid X_{\alpha_2}, ..., X_{\alpha_n})$ is non-random.

Usually $\mathscr{A}=\mathbb{R}_+$ or $\mathbb{N}$ or it is a finite subset of $\mathbb{N}$. In the last case we write $\mathrm{GCS}_2(n)$, where n is a number of elements in $\mathscr{A}$. By Gauss($\mathscr{A}$) we denote a family of the Gaussian random elements fulfilling (i) and by $\mathrm{GCS}_2(n; \mu, \Sigma)$ n-dimensional random vectors from $\mathrm{GCS}_2(n)$ with an expectation μ and a covariance matrix Σ. Obviously for any Gaussian random element the conditions (ii) and (iii) hold. The technical condition (i) is introduced to avoid uninteresting cases of independence of com-

Received October 18, 1990; revised December 10, 1990.

AMS 1980 subject classifications: 62H05, 62E10.

Key words and phrases: Kagan class, Gaussian conditional structure of the second order, multivariate normal law, linearity of regression, homoscedasticity, characterization.

0047-259X/91 $3.00

683/39/1-6

ponents and the situation when a random variable is a linear combination of others—then the conditions (ii) and (iii) are fulfilled trivially.

The random elements with properties (i)–(iii) or very close to them have been intensively investigated in the last years. Here is a comprehensive review of the most interesting results (in some cases technical details are omitted):

(a) $\mathrm{GCS}_2(\mathbb{R}_+) = \mathrm{Gauss}(\mathbb{R}_+)$;

(b) $\mathrm{GCS}_2(\mathbb{N}) = \mathrm{Gauss}(\mathbb{N})$;

(c) $\mathrm{GCS}_2(2) \neq \mathrm{Gauss}(2)$;

(d) $X \in \mathrm{GCS}_2(\mathscr{A}) \Rightarrow \forall p > 0, \forall \alpha \in \mathscr{A}, E\,|X_\alpha|^p < \infty$;

(e) $X, Y \in \mathrm{GCS}_2(n; \mu, \Sigma) \Rightarrow \forall c_1,\ c_2 \geqslant 0,\ c_1 + c_2 = 1,\ c_1 X + c_2 Y \in \mathrm{GCS}_2(n; \mu, \Sigma)$;

(f) $\mathrm{GCS}_2(n) \cap \mathrm{EC}(n) = \mathrm{Gauss}(n)$ ($\mathrm{EC}(n)$ denotes n-variate elliptically contoured measures);

(g) $\mathrm{GCS}_2(n) \cap \mathrm{ID}(n) = \mathrm{Gauss}(n)$ ($\mathrm{ID}(n)$ denotes n-variate infinitely divisible measures).

For further information and details see: Plucińska [10]; Wesołowski [12] for (a); Bryc and Plucińska [5] for (b, c, d); Bryc [4] for (a, d, e); Szabłowski [11] for (f); and Wesołowski [13] for (g).

Similar investigations concerning Gaussian conditional structure were lead by Fisk [6], Brucher [3], Ahsanullah [1], Arnold and Pourahmadi [2]. However, they were interested in stronger assumptions of Gaussian conditional distributions, not in conditional moments only, as we are.

In this note we relate the notion of the $\mathrm{GCS}_2(n)$-distribution to the classification of multivariate probability measures introduced in Kagan [7], where the following definition was proposed.

DEFINITION 2. A random vector $X = (X_1, ..., X_n)$ belongs to the class $\mathscr{D}_{n,k}(\mathrm{loc})$, $k = 1, ..., n$, $n = 1, 2, ...$, if its characteristic function ϕ in a neighbourhood of the origin $V \subset \mathbb{R}^n$ has the form

$$\phi(t) = \prod_{1 \leqslant i_1 < \cdots < i_k \leqslant n} R_{i_1 \cdots i_k}(t_{i_1}, ..., t_{i_k}),$$

$t = (t_1, ..., t_n) \in V$, where $R_{i_1 \cdots i_k}$ is a continuous complex function with $R_{i_1 \cdots i_k}(0) = 1$, $1 \leqslant i_1 < \cdots < i_k \leqslant n$.

The notion is a generalization of the concept of independence of random variables. In Kagan [7] many interesting properties and some useful applications are given. They include extensions of the Darmois–Skitovitch theorem. The investigations are continued in Kagan [8] and Wesołowski [14].

In this paper we consider the class $\mathscr{D}_{n,2}(\text{loc})$ since $\text{Gauss}(n) \subset \mathscr{D}_{n,2}(\text{loc}) \subset \mathscr{D}_{n,k}(\text{loc})$ for any $k \geqslant 2$—see Property (v) in Kagan [7].

In Section 2 we give our main result together with some comments and auxiliary lemmas. All the proofs are put off to Section 3.

2. A Characterization of the Multivariate Normal Law

It is conjunctured that $\text{Gauss}(n) \neq \text{GCS}_2(n)$. (However, the only counter-example known is in the case $n=2$—see (c) in Section 1.) Thus a natural way of investigation is to relate $\text{GCS}_2(n)$ distributions to the known families of probability measures. Two consequences of such a method are the results (f) and (g) from Section 1. Our characterization is another contribution in this direction.

THEOREM. *If $n \geqslant 3$ then*

$$\text{GCS}_2(n) \cap \mathscr{D}_{n,2}(\text{loc}) = \text{Gauss}(n).$$

The result being the core of the proof seems to be of independent interest.

We consider 3-variate random vector X with $\psi(t) = \ln(E\{\exp(i(t, X))\})$ in some neighbourhood of the origin $V \subset \mathbb{R}^3$. Define for $t = (t_1, t_2, t_3) \in V$,

$$H(t) = \psi(t) - \psi(t_1, t_2, 0) - \psi(t_1, 0, t_3) - \psi(0, t_2, t_9).$$

PROPOSITION. *If $X \in \text{GCS}_2(3)$ and*

$$\left.\frac{\partial H(t)}{\partial t_i}\right|_{t_i=0} = \text{const}, \tag{1}$$

$$\left.\frac{\partial^2 H(t)}{\partial t_i^2}\right|_{t_i=0} = \text{const}, \tag{2}$$

$i = 1, 2, 3$, *then all bivariate marginal distributions of X are Gaussian.*

Let us observe that the function H has a very simple form for $\mathscr{D}_{3,2}(\text{loc})$ random vectors. It is a consequence of the following:

LEMMA 1. *If $X \in \mathscr{D}_{3,2}(\text{loc})$ then in a neighbourhood of the origin $V \subset \mathbb{R}^3$ a characteristic function of the X has a form*

$$\phi(t) = \frac{\phi(t_1, t_2, 0)\, \phi(t_1, 0, t_3)\, \phi(0, t_2, t_3)}{\phi(t_1, 0, 0)\, \phi(0, t_2, 0)\, \phi(0, 0, t_3)}.$$

The general version of this result was obtained in Wesołowski [14]. However, the proof of our special case is quite short and to make the paper more self-contained we give it in Section 3.

For the proof of the next lemma we refer to Kagan [7].

LEMMA 2. *If* $X \in \mathscr{D}_{n,k}(\mathrm{loc})$ *and all k-dimensional marginal distributions are Gaussian then X has a n-variate Gaussian distribution.*

Let us observe that from the proof of our main result (see Section 3) it follows that it suffices to assume that only all trivariate marginals belong to $\mathrm{GCS}_2(3)$ (obviously each of them is a $\mathscr{D}_{3,2}(\mathrm{loc})$ random vector). We expect that the following extension of our theorem is true: For any $k = 2, ..., n-1$, $n \geq 3$, $\mathrm{GCS}_2(n) \cap \mathscr{D}_{n,k}(\mathrm{loc}) = \mathrm{Gauss}(n)$. However, an adaptation of our method in the general case does not seem to be simple.

3. PROOFS

Proof of the theorem. If $Y = (Y_1, ..., Y_n)$ belongs to $\mathrm{GCS}_2(n) \cap \mathscr{D}_{n,2}(\mathrm{loc})$ then for any trivariate marginal $(Y_i, Y_j, Y_k) = (X_1, X_2, X_3) = X$, $i, j, k = 1, ..., n$, $i \neq j \neq k \neq i$, we have $X \in \mathrm{GCS}_2(3) \cap \mathscr{D}_{3,2}(\mathrm{loc})$—see Property (II) in Kagan [7]. From Lemma 1 we obtain

$$H(t) = -\psi(t_1, 0, 0) - \psi(0, t_2, 0) - \psi(0, 0, t_3).$$

Consequently (1) and (2) hold. By Proposition we conclude that any bivariate marginal distribution of the Y is normal. Now the result follows from Lemma 2 since $Y \in \mathscr{D}_{n,2}(\mathrm{loc})$. ∎

Proof of the proposition. Without any loss of generality we can assume: $EX_i = 0$, $EX_i^2 = 1$, $i = 1, 2, 3$. Then from the assumption $X \in \mathrm{GCS}_2(3)$ we have in V,

$$\left.\frac{\partial\psi(t_i, t_j, 0)}{\partial t_i}\right|_{t_i=0} = \rho_{ij} \frac{\partial\psi(0, t_j, 0)}{\partial t_j}, \tag{3}$$

$$\left.\frac{\partial\psi(t_i, t_j, t_k)}{\partial t_i}\right|_{t_i=0} = a_{j(i|j,k)} \frac{\partial\psi(0, t_j, t_k)}{\partial t_j} + a_{k(i|j,k)} \frac{\partial\psi(0, t_j, t_k)}{\partial t_k}, \tag{4}$$

$$\left.\frac{\partial^2\psi(t_i, t_j, 0)}{\partial t_i^2}\right|_{t_i=0} = \rho_{ij}^2 - 1 + \rho_{ij}^2 \frac{\partial^2\psi(0, t_j, 0)}{\partial t_j^2}, \tag{5}$$

$$\left.\frac{\partial^2\psi(t_i, t_j, t_k)}{\partial t_i^2}\right|_{t_i=0} = -b_{(i|j,k)} + a_{j(i|j,k)}^2 \frac{\partial^2\psi(0, t_j, t_k)}{\partial t_j^2}$$
$$+ 2a_{j(i|j,k)}a_{k(i|j,k)} \frac{\partial^2\psi(0, t_j, t_k)}{\partial t_j \partial t_k}$$
$$+ a_{k(i|j,k)}^2 \frac{\partial^2\psi(0, t_j, t_k)}{\partial t_k^2}, \tag{6}$$

where $\psi(t_i, t_j, t_k) = \ln\{E\exp[i(t_iX_i + t_jX_j + t_kX_k)]\}$, $i \neq j \neq k \neq i$. It is an immediate consequence of interpretation of the assumptions on conditional moments in terms of characteristic functions given in Kagan *et al.* [9, Lemma 1.1.3]. The coefficients involved in (3)–(6) have the form: $\rho_{ij} = EX_iX_j$ is a correlation coefficient of X_i and X_j,

$$a_{j(i|j,k)} = \frac{\rho_{ij} - \rho_{ik}\rho_{jk}}{1-\rho_{jk}^2}, \qquad a_{k(i|j,k)} = \frac{\rho_{ik} - \rho_{ij}\rho_{jk}}{1-\rho_{jk}^2},$$
$$b_{(i|j,k)} = |K|/(1-\rho_{jk}^2),$$

where $|K|$ is the determinant of the covariance matrix K of the X. Obviously from the assumptions we have

$$0 < |\rho_{ij}| < 1, \qquad i, j = 1, 2, 3, i \neq j. \tag{7}$$

We denote $a = a_{j(i|j,k)}$ and $b = a_{k(i|j,k)}$. Observe that

$$\rho_{ij} = a + \rho_{ik}b, \qquad \rho_{ik} = b + \rho_{jk}a. \tag{8}$$

We apply (1) and (2) to (3)–(6) and take into account (8). After some computing we obtain the following partial differential equation of the second order

$$ab\frac{\partial^2\psi(0, t_j, t_k)}{\partial t_j^2} + (a^2+b^2)\frac{\partial^2\psi(0, t_j, t_k)}{\partial t_j \partial t_k}$$
$$+ ab\frac{\partial^2\psi(0, t_j, t_k)}{\partial t_k^2} + a\rho_{ik} + b\rho_{ij} = 0. \tag{9}$$

It is a hyperbolic equation in the case $|a| \neq |b|$ and a parabolic equation for $|a| = |b|$. (Observe that by (7) the case $a = b = 0$ is impossible.) If $|a| \neq |b|$ then the general solution of (9) has a form

$$\psi(0, t_j, t_k) = f(at_j - bt_k) + g(bt_j - at_k) - \frac{a\rho_{ik} + b\rho_{ij}}{a^2+b^2} t_jt_k,$$

where f and g are C_2-functions. To obtain the exact solution we first apply (3) to the above formula. Thus

$$\rho_{ij} f'(bt) + \rho_{ik}\, g'(at) = -t(a\rho_{ik} + b\rho_{ij})/(a^2 + b^2).$$

We differentiate the equation with respect to t and obtain

$$b\rho_{ij} f''(bt) + a\rho_{ik}\, g''(at) = -(a\rho_{ik} + b\rho_{ij})/(a^2 + b^2). \tag{10}$$

Second, (5) together with our general solution yields

$$\rho_{ij}(a - b\rho_{jk})\, f''(bt) + \rho_{ik}(b - a\rho_{jk})\, g''(at) = \rho_{jk}^2 + 1. \tag{11}$$

Since the determinant of the system of linear equations (10) and (11) (with $f''(bt)$ and $g''(at)$ unknown) is $\rho_{ij}\rho_{ik}(a^2 - b^2) \neq 0$, then $f''(bt)$ and $g''(at)$ do not depend on t. Hence f and g are quadratic functions and thus in a neighbourhood of the origin the characteristic function of any bivariate marginal distribution has a form $\exp[Q(s, t)]$, where Q is a quadratic form. Consequently it is a characteristic function of a bivariate Gaussian law.

Now consider the case $a = \pm b$. Then $\rho_{ik} = \pm\rho_{ij}$ and (9) takes the form

$$\frac{\partial^2\psi(0, t_j, t_k)}{\partial t_j^2} \pm 2\,\frac{\partial^2\psi(0, t_j, t_k)}{\partial t_j\,\partial t_k} + \frac{\partial^2\psi(0, t_j, t_k)}{\partial t_k^2} = -2(1 \pm \rho_{jk}).$$

The general solution of the above parabolic equation has the form

$$\psi(0, t_j, t_k) = -(1 \pm \rho_{jk})\, t_k^2 + t_k f(t_k \mp t_j) + g(t_k \mp t_j),$$

where f and g are C_2-functions.

Now we apply (3) to our general solution and obtain

$$t_k f'(t_k) + g'(t_k) = \pm 2\rho_{jk} \mp \frac{\rho_{jk}}{1 \pm \rho_{jk}} f(t_k)$$

and, after differentiation,

$$t_k f''(t_k) + g''(t_k) = \pm 2\rho_{jk} \mp \frac{2\rho_{jk} \pm 1}{1 \pm \rho_{jk}} f'(t_k).$$

On the other hand, by (5)

$$t_k f''(t_k) + g''(t_k) = -1 - \frac{2\rho_{jk}^2}{1 \mp \rho_{jk}} + \frac{2\rho_{jk}^2}{1 \mp \rho_{jk}^2} f'(t_k).$$

Both the above equations imply $f' = 1 - \rho_{jk}$. Hence f is a linear function. Thus by any of the equations g'' is constant and consequently g is a quadratic function. Now the form of the general solution ψ yields the final conclusion. ∎

Proof of Lemma 1. If ϕ is a characteristic function of a $\mathscr{D}_{3,2}$(loc)-random vector then from the definition

$$\phi(t_1, t_2, t_3) = R_{12}(t_1, t_2)\, R_{23}(t_2, t_3)\, R_{13}(t_1, t_3) \tag{12}$$

in a neighbourhood of zero. From (12) we have

$$\begin{aligned}
\phi(t_1, t_2, 0) &= R_{12}(t_1, t_2)\, R_{23}(t_2, 0)\, R_{13}(t_1, 0),\\
\phi(t_1, 0, 0) &= R_{12}(t_1, 0)\, R_{13}(t_1, 0),\\
\phi(0, t_2, 0) &= R_{12}(0, t_2)\, R_{23}(t_2, 0).
\end{aligned}$$

Consequently,

$$R_{12}(t_1, t_2) = \frac{\phi(t_1, t_2, 0)\, R_{12}(t_1, 0)\, R_{12}(0, t_2)}{\phi(t_1, 0, 0)\, \phi(0, t_2, 0)}.$$

Similar reasoning for R_{23} and R_{13} gives the result since also by (12) we have

$$\begin{aligned}
&\phi(t_1, 0, 0)\, \phi(0, t_2, 0)\, \phi(0, 0, t_3)\\
&\quad = R_{12}(t_1, 0)\, R_{12}(0, t_2)\, R_{23}(t_2, 0)\, R_{23}(0, t_3)\, R_{13}(t_1, 0)\, R_{13}(0, t_3). \quad ∎
\end{aligned}$$

ACKNOWLEDGMENT

The author is greatly indebted to the referee for a suggestion to take a closer look at the case $|a| = |b|$ in Eq. (9). Its treatment in the first draft of the paper appears to be incorrect.

REFERENCES

[1] AHSANULLAH, M. (1985). Some characterizations of the bivariate normal distribution. *Metrika* **32** 215–218.

[2] ARNOLD, B. C., AND POURAHMADI, M. (1988). Conditional characterizations of multivariate distributions. *Metrika* **35** 99–108.

[3] BRUCHER, J. (1979). A note on the bivariate normal distribution. *Comm. Statist.* **8** 175–177.

[4] BRYC, W. (1985). Some remarks on random vectors with nice enough behaviour of conditional moments. *Bull. Polish Acad. Sci. Math.* **33** 677–684.

[5] BRYC, W., AND PLUCIŃSKA, A. (1985). A characterization of infinite Gaussian sequences by conditional moments. *Sankhyā Ser. A* **47** 166–173.

[6] Fisk, P. R. (1970). A note on a characterization of the multivariate normal distribution. *Ann. Math. Statist.* **41** 486–494.

[7] Kagan, A. M. (1988). New classes of dependent random variables and a generalization of the Darmois–Skitovitch theorem to the case of several forms. *Teor. Veroyatnost. Primenen.* **33** 305–315. [Russian]

[8] Kagan, A. M. (1989). Generalized condition of identity of distributions of random vectors in connection with the analytic theory of linear forms in independent random variables. *Teor. Veroyatnost. Primenen.* **34** 370–375. [Russian]

[9] Kagan, A. M., Linnik, J. V., and Rao, C. R. (1973). *Characterization Problems of Mathematical Statistics.* Wiley, New York.

[10] Plucińska, A. (1983). On a stochastic process determined by the conditional expectation and the conditional variance. *Stochastics* **10** 115–129.

[11] Szabłowski, P. J. (1989). Can the first two conditional moments identify a mean square differentiable process? *Comput. Math. Appl.* **18** 329–348.

[12] Wesołowski, J. (1984). A characterization of the Gaussian process based on properties of conditional moments. *Demonstratio Math.* **18** 795–808.

[13] Wesołowski, J. (1990). *Multivariate Infinitely Divisible Distributions with Gaussian Second Order Conditional Structure.* Math. Inst. Warsaw Univ. of Tech. Preprint No. 17, pp. 1–6.

[14] Wesołowski, J. (1991). On determination of multivariate distribution by its marginals for the Kagan class. *J. Multivariate Anal.* **36** 314–319.